

UNCOVERING THE RELATIONSHIP BETWEEN CATERING SERVICE LOCATION AND URBAN SPATIAL STRUCTURE BASED ON SNS DATA

○LI Miaoyi*¹ SHEN Zhenjiang*²
KOBAYASHI Fumihiko*³
NISHINO Tatsuya*⁴

Keywords : Dazhongdianping, Big data, Restaurant distribution, Space syntax

1. INTRODUCTION

Today, informational applications with advanced technology have great impact on people's daily life in metropolises. Since 1990s, one popular issue in the fields of urbanism and architecture is how informational technologies can affect the locational logic of commercial structure and shopping behavior in urban places. With the rise of 'global market' and further intensification of economical networks supported by these informational technologies, some scholars predicted that the proximity of places will eventually become no relevance with the demise of geographical space' (Manuel, 1996).

However, this pessimistic viewpoint is becoming more and more contested. At the regional or global scale, they believe that in the information age, the centrifugal and centripetal forces co-existed as two sides of a coin for urban spaces in the world (Liu, 2002; Gillespie et al., 2000). On the one hand, the advance of informational technologies leads to dispersion of manufactory activities towards marginal places. On the other hand, management and services are becoming more and more concentrated in few global cities (Sassen, 2001).

For some experience-based commercial functions such as restaurants in urban area, the online browsing and reviewing also provide new possibilities for both customers and restaurant owners. On some of these very popular websites, such as Dazhongdianping in China, the user can easily select the restaurants based on several filters like the location, cost, popularities and ranks evaluated by customers. Together with the use of GPS system in cars and mobile phones, it seems location of restaurant does not matter to customers, at least not as much as before. However, some scholars hold different viewpoint. Some researchers pointed out that when the informational accessibility is no longer a problem, the role played by actual spatial accessibility would rather be enhanced (Read, 2009).

In this line of thinking, we believe that the value of real

place experience is even more valuable and influential than before and thus the spatial accessibility will be enhanced by recognition via the Internet. Today, the huge amount of data on these information platforms opens up new possibilities for conventional urban models. For the space syntax research, it can reduce the fieldwork and help to choosing the appropriate spatial parameters. In present research we will use the big data by information platform to prove this viewpoint.

2. MATERIALS AND METHODS

2.1. About Dazhongdianping data

Dazhongdianping Company is founded in 2003 in Shanghai. Its website (www.dianping.com) is one of the first Chinese web forum and service platform which provides information for the convenience of people's daily life. During 10 years development, Dazhongdianping Website has become a leading one in the field of service, and to a large extent it changes the way people using catering functions in cities. Users can make their own choices based on many characters such as the style of taste, quality of environment and service, the average cost, etc. It also shows the reviews and comments from previous customers and the average rank of certain restaurant based on the customer's feedbacks.

The data collection of this research started from May 2013. The research area covers 5 main administrative districts of Beijing city, China, which are respectively Dongcheng, Xicheng, Chaoyang, Haidian and Fengtai. Within each district, there are 10 to 20 smaller sub-areas listed on the website.

In the default setting of the website, each page can show up to 15 restaurants in one page, practically the restaurant listed in the first page are the most visible ones. Thus, for each sub-area, we select top 15 restaurants in the default-ranking filter. After delete the number of some restaurants reputedly calculated in sub-areas, there are in total 1123 restaurants selected as samples. For each restaurant, three types of data are recorded including the average cost per meal per person (in RMB), the number of reviews from previous customers and the average rank based on their

Table-1 The relationship between reviewing number, average cost and rank with spatial parameters

Review	NodeCount	NACH1km	NACH2km	NACH5km	NACH10km	NACH20km	NACH50km	INT1km	INT2km	INT5km	INT10km	INT20km	INT50km
		0.89211	0.89138	0.87704	0.863679	0.848731	0.820673	69.3338	209.023	1049.41	3530.9	9851.6	16258.2
top20%review	77	1.10456	1.18803	1.22583	1.22733	1.21205	1.1974	134.976	442.461	2136.34	6889.58	16857.8	21674.8
		123.81%	133.28%	139.77%	142.10%	142.81%	145.90%	194.68%	211.68%	203.58%	195.12%	171.12%	133.32%
top20~40%review	219	1.16499	1.21273	1.22063	1.20463	1.18478	1.17624	131.28	416.673	2010.06	6351.53	15713.5	21087
		130.59%	136.05%	139.18%	139.48%	139.59%	143.33%	189.34%	199.34%	191.54%	179.88%	159.50%	129.70%
top40~60%review	316	1.13333	1.18467	1.18753	1.17254	1.14927	1.13302	121.907	391.501	1898.49	6051.23	15175.9	20612.6
		127.04%	132.90%	135.40%	135.76%	135.41%	138.06%	175.83%	187.30%	180.91%	171.38%	154.05%	126.78%
top60~80%review	147	1.10623	1.15365	1.15105	1.12795	1.10383	1.08311	108.515	338.519	1603.55	5268.17	13766.9	19648.9
		124.00%	129.42%	131.24%	130.60%	130.06%	131.98%	156.51%	161.95%	152.80%	149.20%	139.74%	120.86%
bottom20%review	53	1.06994	1.11407	1.10267	1.08068	1.06247	1.04644	95.9451	274.991	1322.26	4429.57	12220.8	18746.7
		119.93%	124.98%	125.73%	125.13%	125.18%	127.51%	138.38%	131.56%	126.00%	125.45%	124.05%	115.31%
Cost	NodeCount	NACH1km	NACH2km	NACH5km	NACH10km	NACH20km	NACH50km	INT1km	INT2km	INT5km	INT10km	INT20km	INT50km
top20%cost	87	0.89211	0.89138	0.87704	0.863679	0.848731	0.820673	69.3338	209.023	1049.41	3530.9	9851.6	16258.2
		1.13807	1.18952	1.19334	1.17523	1.15451	1.13932	121.086	395.721	1922.75	6248.63	15558	20825.4
top20~40%cost	190	1.13084	1.17367	1.17241	1.15312	1.13143	1.11359	123.689	388.308	1860.93	6048.71	15183.3	20578.2
		126.76%	131.67%	133.68%	133.51%	133.31%	135.69%	178.40%	185.77%	177.33%	171.31%	154.12%	126.57%
top40~60%cost	286	1.11632	1.17745	1.18954	1.17822	1.15764	1.14475	121.557	384.625	1864.44	5855.76	14770.2	20599.3
		125.13%	132.09%	135.63%	136.49%	136.40%	139.49%	175.32%	184.01%	177.67%	165.84%	149.93%	126.70%
top60~80%cost	122	1.11255	1.1637	1.16416	1.14765	1.12522	1.10806	108.799	349.417	1749.42	5632.29	14422.6	20063.3
		124.71%	130.55%	132.74%	132.88%	132.58%	135.02%	156.92%	167.17%	166.71%	159.51%	146.40%	123.40%
bottom20%cost	53	1.19496	1.22514	1.21626	1.19091	1.16778	1.15707	125.159	388.924	1808.02	5873.82	14963.6	20101.7
		133.95%	137.44%	138.68%	137.89%	137.59%	140.99%	180.52%	186.07%	172.29%	166.35%	151.89%	123.64%
Rank	NodeCount	NACH1km	NACH2km	NACH5km	NACH10km	NACH20km	NACH50km	INT1km	INT2km	INT5km	INT10km	INT20km	INT50km
top20%	233	0.89211	0.89138	0.87704	0.863679	0.848731	0.820673	69.3338	209.023	1049.41	3530.9	9851.6	16258.2
		1.14196	1.18134	1.1774	1.15875	1.13108	1.11352	128.573	409.383	1995.86	6550.92	16218.1	20713.2
top20~40%	227	1.14723	1.19838	1.2092	1.19603	1.17791	1.16652	126.005	408.463	1984.12	6271.66	15517.3	20982.8
		128.60%	134.44%	137.87%	138.48	138.78%	142.14%	181.74%	195.42%	189.07%	177.62%	157.51%	129.06%
top40~60%	286	1.14353	1.19618	1.20021	1.18027	1.15899	1.14578	115.71	359.569	1700.64	5373.02	13916.6	20127.2
		128.18%	134.19%	136.85%	136.66%	136.56%	139.61%	166.89%	172.02%	162.06%	152.17%	141.26%	123.80%
top60~80%	52	1.00467	1.08023	1.07827	1.07521	1.05859	1.0449	100.15	316.594	1604.39	5142.9	13388.4	19940.7
		112.62%	121.19%	122.94%	124.49%	124.73%	127.32%	144.45%	151.46%	152.88%	145.65%	135.90%	122.65%
bottom20%	39	1.05988	1.13	1.14442	1.12318	1.10549	1.07509	102.563	300.742	1364.96	4459.79	12274.8	19111.5
		118.81%	126.90%	130.49%	130.05%	130.25%	131.00%	147.93%	143.88%	130.07%	126.31%	124.60%	117.55%

reviews (0-5 stars). Among these data, the number of reviews could be comparable to the number of visitors between these restaurants.

To get comments data from website, a web crawler is designed to obtain data automatically. In general, the data acquisition process includes four steps: get and parse the web page, geocoding using the name or address attribute, store the results into database and generate shapefile file using database table. In order to obtain comments data as soon as possible, we use the python language and some third-party libraries, such as the requests package to get web page, the pyquery package to parse web page to get interested information. To geocoding the data and get the longitude and latitude of each restaurant, we use an open and free web service named Baidu Geocoding API. This service is stable and thus can always return credible results. To store the results, open source database MySQL is used. As well, to generate shapefile in order to analysis, an efficient python package named pyshp is utilized.

2.2. Integration and Choice Value in Space Syntax

The area of space syntax analysis covers the whole metropolitan region of Beijing city, including its satellite towns such as Changing, Shunyi, Tongzhou and Mentougou. The analysis is mainly focusing on two types of spatial parameters in the angular analysis mode: integration and choice value.

Integration is an old spatial parameter could be traced to the origin of space syntax tool. It measures the topological distances of end points of each street segment to other points

of road segments in certain range. The choice value (normalized angular choice) used in this research is a new development of space syntax in 2012 (Hillier et al., 2012). Based on the existed empirical study of 50 cities (Hillier et al., 2012), this measurement is proved to strongly related with the automobile traffic.

The choice can show different scales of movement networks. In terms of shopping behavior, integration and choice value can reveal two different kinds of spatial potentials: integrations show the

potential of a place (or a shop inside a shopping mall) being visited as a target of the trip. While choice shows the potential of a place being passing-by by chance. Using these two parameters, this research will explore how they can influence the shopping behaviors in different scales.

3. URBAN-SCALE ANALYSES ON DAZHONGDIANPING DATA

This research starts with analyzing the relationship of three sets of data from Dazhongdianping website (reviewing number, average cost and rank) with the spatial parameters. Table-1 lists their locations' integration and choice value in different radius. The darker the background color means the higher dependency of certain group of data on certain type of spatial parameters. For instance, in the list of reviewing number, 77 restaurants with top 20% number of reviews tend to be influenced by 50km radius choice value most. Their location's choice value in 50km radius is above the average rate by 45%.

As the Table-1 shows, the reviewing number is more influenced by spatial parameters than rank and average cost. The top 20% reviewed restaurants tend to locate on the street segments with high 50km radius choice value and high 2km radius integration value. In fact, the table shows that other levels of reviewing numbers reveal the similar tendency.

It is necessary to point out that the regression analysis on three sets of data (reviewing number, average cost and rank) with the spatial parameters leads to a very low correlation.

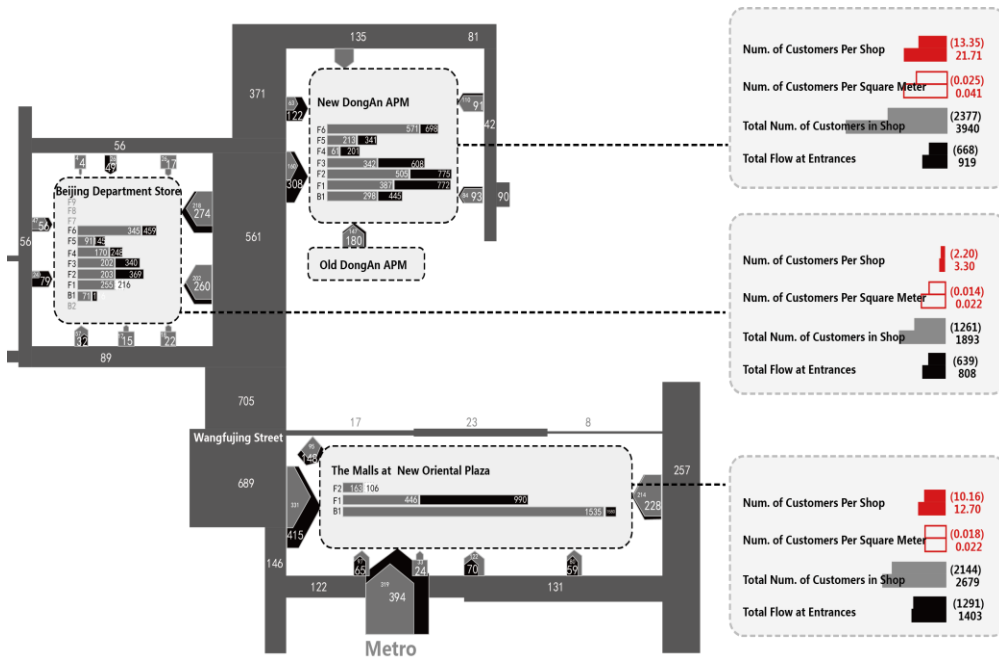


Figure-1. The flow intensity in Wangfujing area and number of customers in three shopping malls (week day and weekend)

The highest R-square value turns to be the relationship between reviewing number and 5km integration value (only 0.131).

One of the possible reasons is the characteristic of catering function. Later research at architecture scale shows that they tend to locate in relatively low visible space comparing with retail functions. Therefore, instead of analyzing the spatial condition of the directly connected street segments, this research takes the local area as a group of streets. Furthermore, because the previous part of present research has indicated the influences of 2km integration and large scale normalized angular choice, this part focuses on the total number of restaurants listed in each small district in Dazhongdianping and the spatial condition of that particular area.

Using the tool of regression analysis with two factors (2km integration and 50km Choice), the R-square value of distribution density of the restaurants on Dazhongdianping and integrated spatial parameter can reach 0.51. The weight between Choice and 2km integration is 4.613:1.

Firstly, the above findings indicate that in the information age the spatial condition still plays important role in the distribution of catering functions. Secondly, this spatial condition could be understood as an integration of two main aspects: how local area is connected in the whole city (50km Choice) and how it is connected in local street fabric itself (2km integration). Furthermore, the connection in city as a whole is more important than its local connections. All of these finding indicate the restaurant which is better connected

in the whole city and local street fabrics can benefit more from Dazhongdianping. They tend to be visited more often and get more reviews.

4. BLOCK SCALE AND ARCHITECTURE SCALE ANALYSIS: FIELDWORK RESEARCH IN WANGFUJING AREA AND THE MALLS AT NEW ORIENTAL PLAZA

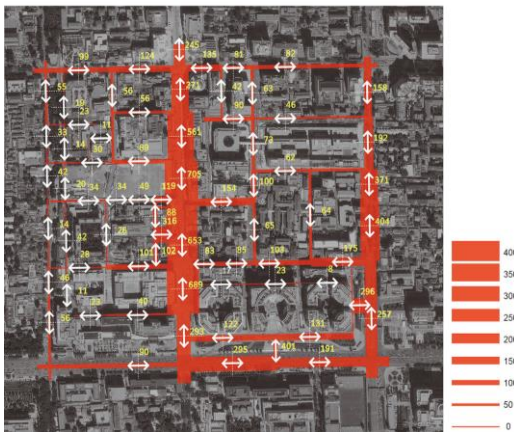
4.1 Data Collection and Fieldwork

Wangfujing area has the highest number of restaurants on Dazhongdianping Website. In this part, top 60 restaurants with highest reviewing numbers are selected as samples from major shopping malls in Wangfujing high street, which are respectively Malls at New Oriental Plaza, Beijing Department Store and New Dong'an APM.

20 restaurants of the top 60 restaurants locate at the Malls at New Oriental Plaza, and 10 locate at the New Dong'an APM, and only 5 in Beijing Department Store. The total number of reviews and estimated profit (average cost multiply the reviewing number) of these three shopping malls are in the same order, but the New Dong'an APM rank first in average number of reviews per restaurants and average cost. The rest 25 restaurants are distributed in the street segments in the whole area. But only 3 of them are located directly on Wangfujing high streets. From this point, catering functions tend to stay away from the main street comparing with other retail functions.

A fieldwork is organized to compare the web data with the actual use of space on the streets and inside the buildings. There are 72 observation points on the streets. Pedestrian flows are measure four times in a day. The fieldwork inside shopping malls was done in two days (one weekday and one weekend day). There are 346 observation points inside the three shopping malls, the flow passing those points were also measured four times a day. The instant number of customers inside 893 shops or restaurants inside 3 case shopping malls was noted down three times a day (morning, lunch time and

Total Pedestrian Flow



Segment Analysis (integration 3km)



A Beijing Department Store	B New Dong An APM	C New Oriental Mall
(Average Choice) : 0.799	(Average Choice): 0.831	(Average Choice) : 0.880
(Average Integration) : 736.0	(Average Integration): 819.254	(Average Integration) : 849.8

Figure-2. Pedestrian flow intensity in Wangfujing area and the spatial analysis of 3km integration value

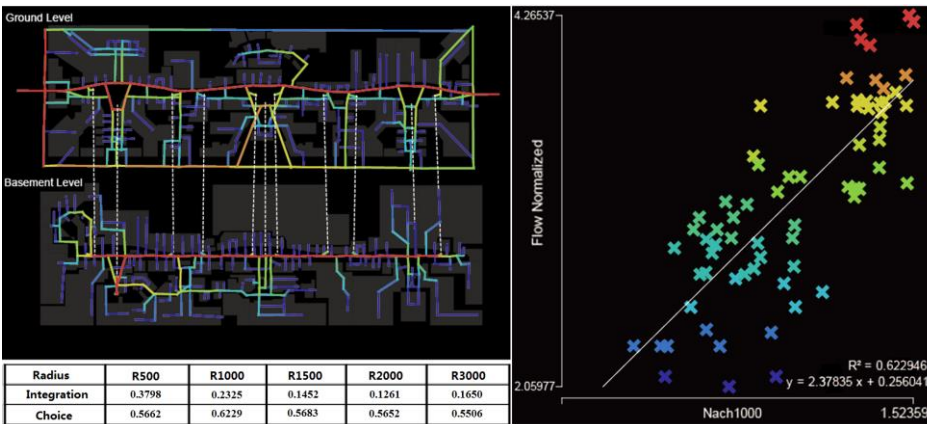


Figure-3. The 3D spatial analysis of the Malls at New Oriental Plaza

including the ground level of major shopping malls is mapped based on the fieldwork. Figure-2 shows the observed pedestrian flow intensity with the space syntax analysis result of 3km integration. The R-square value is 0.612. Furthermore, the average 3km integration and choice value of three major shopping malls are also listed.

The result shows the Malls at New Oriental Plaza has the highest integration and choice value while the Beijing Department Store ranks in the last. This ranking fits with the statistical data from Dazhongdianping Website and the fieldwork very well. A detail study on the case of New Oriental Plaza starts by using the ‘unlink’ tool in space syntax, a 3D model of the mall is established (see Figure-3). 1km choice value is strongly correlated with the observed

afternoon).

Figure-1 visualizes the flow intensity on the streets and at the entrances of 3 shopping malls, it also illustrates the number of customers inside shops in each floor in weekday and weekend. The flow at the entrances shows the Malls at New Oriental Plaza has highest number of people coming in and out of the building, while Beijing Department Store has lowest number. Comparing the number in weekday and weekend, the malls at the New Oriental Plaza tends to run a very stable business while the New Dong’an APM tends to benefit most from the weekend or holiday.

4.2 Spatial analysis on Wangfujing of shopping mall and entertainment area

This part will use space syntax tool to analyze the spatial parameters of the Wangfujing area and the case shopping malls. It will compare the actual spatial condition and usage with the virtual visibility. Furthermore, a more detail study in the Malls at New Oriental Plaza will be presented.

At the urban block scale, a detailed pedestrian network

flow inside the building (R-square=0.623).

For experience-based shops such as restaurants, the result suggests they mostly gather on less visible or blind alley. But among these spaces those located in relative more visible space and those have larger area shops can attract more customers. We integrate the data of average cost with the actual number of customers in these restaurants, the result shows an even more clear spatial logic (see Figure-4). Highly visible shops tend to have more customer with proper cost (variety of different price choice as well). These shops can guarantee the building a stable business and higher ability to pay the rent in long term.

When focusing on the catering functions inside the building, most of them are located on the least visible and least accessible part of corridors with the Macdonald’s as the only exception. They locate on the underground floor, at the east and west end of the building and a sideway in the south part which is parallel with the main corridor. This distribution

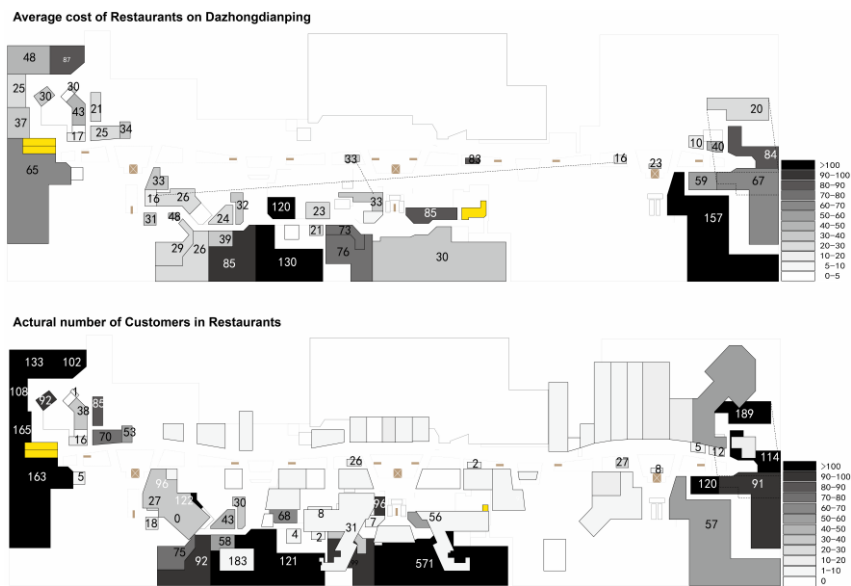


Figure-4 The comparison of actual number of customers, average cost on web

is both a design strategy and a natural logic for the catering function because they are only visited in certain times of a day. Similar distribution pattern can also be found in previous study in block scale.

5. CONCLUSION

It seems that any kind of claims underestimated the impact of the information technologies on our contemporary urban life is so powerless. Indeed, without Dazhongdianping, it is very hard to get the similar set of data in this research. Especially at urban scale it will be a mission impossible.

Dazhongdianping greatly changed our ways of eating out in cities. With the dissolving of the information barriers, it seems that the role of geological space should be weakened. However, catering functions or some other type of commercial functions like theater still depends on the physical experience in actual space. For this type of experience-based functions, the location still matters. A good location can guarantee more customers, and tend to be reviewed more often on web.

At the urban block scale, a detailed pedestrian network including the ground level of major shopping malls is mapped based on the fieldwork. Figure-2 shows the observed pedestrian flow intensity with the space syntax analysis result of 3km integration. The R-square value is 0.612. Furthermore, the average 3km integration and choice value of three major

In detail, this research can give us some preliminary empirical findings: at urban scale most popular restaurants on web are located in the area which is well connected in the city as a whole and also topological centered in the local street fabrics. At urban block scale, the well connected shopping

malls tend to have more popular restaurants and therefore the study shows a clear spatial logic. But in the architecture scale there is a clear difference between the observed number of customers and the reviewing number. This result suggests that a target consumption based on web depends less on the spatial configuration inside the building, but more on the location of the building in its urban context. The limitation of this research so far is lack of further empirical data and the improvement of the survey. Noting down the instant number of customers enlarges the influence of shopping area. A better

method is to measure the flow coming in and out of each individual shop.

As a summary, virtual space is a tool to facilitate the shopping behavior by increasing its information accessibility. In this sense, it is just as the transportation technologies, which can increase the physical accessibility of these shops. As tools they are normally being implemented in an already successful area while leave some other less successful area by-passed.

REFERENCES

- 1) Manuel Castells, *The Information Age: Economy, Society and Culture* [M], *The Rise of the Network Society*, Blackwell, 1996.
- 2) Liu Weidong. Development of the Internet in China: spatial characteristics and implications [J]. *Geographical Research*, 2002, 21(3): 347-356.
- 3) Gillespie A, Richardson R, Cornford J. *Regional Development and the New Economy*, Research paper of the Center for Urban and Regional Development Studies [J], University of Newcastle Upon Tyne, UK, 2000.
- 4) Sassen. S., *The Global City: New York, London, Tokyo*. [M], Princeton: Princeton University Press, 2001.
- 5) Read, S. Another Form: From the 'Informational' to the 'Infrastructural' City [J], in *Footprint*, 2009, 5: 5-22.
- 6) Hillier, B., Yang, T., Turner, A., Advancing DepthMap to advance our understanding of cities: comparing streets and cities, and streets to cities[C]. in *Eighth International Space Syntax Symposium*. 2012. Santiago de Chile: PUC.

*1 Doctoral Course, Graduate School of Natural Science & Technology, Kanazawa University.

*2 Professor, Graduate School of Natural Science & Technology, Kanazawa University, Dr. Eng.

*3 Lecturer, Graduate School of Natural Science & Technology, Kanazawa University, Dr. Eng.

*4 Associate Professor, Graduate School of Natural Science & Technology, Kanazawa University, Dr. Eng.

SNS を用いた飲食店の立地と都市空間構造の関連性の把握に関する研究

○李 苗裔*¹ 沈 振江*²
小林 史彦*³ 西野 辰哉*⁴

キーワード: Dazhongdianping, ビッグデータ, 飲食店分布, スペースシタクス

概要:

情報社会の発達は日常生活に大きな影響を与えるようになった。1990年代以来、情報技術の普及によって、都市空間に立地する飲食店や商店の場所も同じように影響を受けることは研究の課題となり、インターネットの情報認知によって、実際の都市空間にある店舗のアクセス性も変わるように思われる。

本研究は北京を取り上げ、SNS のデータを用いて主に飲食店の人気度と立地の関連を明らかにすることを目的としている。本文で取り上げる「大衆点評」の SNS サイトは、中国に最初に関連サービスを提供したサイトで、都市居住者が自ら日常生活を記録することができる。SNS の利用者は自分の好みで、飲食店の雰囲気やサービスを評価することができ、コメントを公開することも自由にできる。また、SNS では、利用者による飲食店のランキングも公開している。本研究では、当該サイトに公開している 2013 年 5 月の時点からのデータを収集し、研究対象地域は、東城、西城、朝陽、海淀と豊台の五つの区からそれぞれ 10-20 の地区を対象にしている。SNS でコメントされた飲食店の空間分布を分析し、さらに現地調査により王府井地区の分布を確認した。

本研究では、街区レベルでは、大型百貨店のまわりに飲食店が多いことを明らかにした。そして、建物の中にある飲食店は、SNS では数が少ないことも判明した。実際の現地調査が不十分であるが、実際の利用状況に比べ、ネット上のコメントは飲食店の影響を拡大する傾向があることは判明した。このため、ネットにある仮想社会は、情報のアクセス性によりネット利用者への影響によって飲食店の売りに影響を与えることが考えられる。

本研究では、交通アクセス性ではなく、情報アクセス性の考察を通して、仮想社会は実社会の飲食店の利用者の数には大きな影響があることを明らかにした。

*1 金沢大学大学院自然研究科 博士後期課程

*2 金沢大学大学院自然研究科 教授 博(工学)

*3 金沢大学大学院自然研究科 講師 博(工学)

*4 金沢大学大学院自然研究科 准教授 博(工学)